

Using Airborne Hyperspectral and Satellite Multispectral Data to Quantify Within-Field Spatial Variability

S. Y. Hong* K. A. Sudduth** N. R. Kitchen** H. L. Palm** and W. J. Wiebold**
USDA-ARS, Cropping Systems and Water Quality Research

ABSTRACT

The relationship between hyperspectral and multispectral remotely sensed images and ground-based soil and crop information was investigated for two central Missouri experimental fields in a corn (*Zea mays* L.)-soybean (*Glycine max* L.) rotation. Multiple airborne hyperspectral and IKONOS images were obtained during the 1999 and 2000 growing seasons. Hyperspectral images (HSI) covered 120 bands from 471 to 828 nm with a spatial resolution of 1 m. Multispectral IKONOS images included four bands ranging from blue, green, red, and near-infrared with a spatial resolution of 4 m. Geometric distortion of the pushbroom-type sensor caused by aircraft attitude change during image acquisition was corrected with a rubber sheeting transformation. Within-field data collection included crop yield, soil electrical conductivity (EC_s), and soil chemical properties. Simple correlation, multiple regression, and principal component analysis were used to identify those hyperspectral data most highly related with field measured soil and crop properties. Blue wavelengths were most highly correlated with EC_s measurements. For corn, the early reproductive stage provided the best relationships between final yield data and spectral signatures in both years. For soybeans, yield data were highly correlated with wavelengths in the near infrared region from August images in both 1999 and 2000. Maps estimating soil EC_s and crop yield from hyperspectral and multispectral images were derived.

Keywords : Remote sensing, Precision agriculture, Hyperspectral images, Multispectral images.

1. INTRODUCTION

Precision agriculture, or site-specific crop management (SSCM), is an information-based management-intensive approach to farming. Instead of managing a field as a whole, the philosophy of precision agriculture is to manage individual areas within a field. Understanding the functional relationship of crop yield to other spatial factors, therefore, is a basic need for successful SSCM (Sudduth et al., 1996). Identification and quantification of within-field soil and crop condition (such as crop density differences, differences in soil properties, stress or damage caused by diseases, weeds, and pesticides) is needed to understand their effect on yield. If crop stress indicators can be spatially located, then an observer can visit that area in the field to diagnose the cause.

Image-based remote sensing (RS) is an efficient

way to detect spatial differences in crop and soil conditions within a field. The recent convergence of technological advances in geographic information systems (GIS), global positioning systems (GPS), and automatic control of farm machinery through variable rate technology (VRT) within the precision crop management system have provided an ideal framework for utilizing RS for farm management (Moran, 2000). Remotely sensed data are also useful in helping to define management units. Remote sensing offers the potential for identifying fine scale spatial patterns in soil properties across a field, and optimizing the soil sampling strategy to quantify those patterns (Mulla et al., 2000).

Imaging spectrometry which is known as hyperspectral sensing is defined as the simultaneous acquisition of images in many relatively narrow, contiguous and/or non-contiguous spectral bands throughout the ultraviolet, visible and infrared

* Soil Management Division, NIAST, RDA, Korea

** University of Missouri, Columbia, Missouri 65203, USA.

portions of the spectrum (Jensen, 2000). The value of an imaging spectrometer lies in its ability to provide a high-resolution reflectance spectrum for each picture element in the image. The reflectance spectrum in the region from 0.4–2.5 μm may be used to identify a large range of surface cover materials that cannot be identified with broadband, low-spectral-resolution imaging systems such as the Landsat MSS, TM, or SPOT (Goetz et al., 1985).

Airborne pushbroom scanning provides an effective method for hyperspectral imaging (HSI) with a low cost digital CCD camera (Mao, 2000). However, the data obtained with an aerial pushbroom HSI system suffers from geometric distortions. Some of the distortions are caused by aircraft attitude change during image scanning. When the aircraft attitude changes, the scanner is presented with an off-nadir scene, causing distortion. This problem is especially severe in the in-track direction due to roll of the aircraft. These distortions must be corrected before the image data can be geo-referenced and used for field pattern identification (Yao et al., 2001).

The objective of this study was to explore the relationships between the spectral reflectance signatures and biophysical properties of crops and soils using HSI, and to evaluate the usefulness of HSI for quantifying within-field spatial variability.

2. GROUND DATA COLLECTION AND PROCESSING

Data were collected on two fields (Field 1, 35

ha and Gvillo, 13 ha) located within 3 km of each other near Centralia, in central Missouri. The soils found at these sites are characterized as claypan soils of the Mexico-Pumam association (fine, smectitic, mesic aeric Vertic Aqualfs). Mexico-Pumam soils formed in moderately fine textured loess over a fine textured pedisegment. Surface textures range from a silt loam to a silty clay loam. The subsoil claypan horizon(s) are silty clay loam, silty clay or clay, and commonly contain as much as 50 to 60% montmorillonitic clay. Within each study field, topsoil depth above the claypan ranged from less than 10 cm to greater than 100 cm. Because of the high-clay subsurface horizons, topsoil depth above the claypan is often correlated to spatial variations in crop productivity (Kitchen et al., 1999).

Ground measurements used in this analysis included combine grain yield, soil electrical conductivity (EC_s), and soil chemical properties. Two corn and two soybean crop-years were obtained (Table 1). Gleaner R42 or R62 combines equipped with AgLeader Yield Monitor 2000 yield sensing systems were used to obtain yield data. Data collection and processing techniques were as described by Birrell et al. (1996). To remove outliers, data points four standard deviations above or below the mean yield were removed, as were data collected at harvesting speeds of less than 0.75 m/s. Yield data were analyzed using geostatistics, and mapped by block kriging (1 m cell size) with appropriate semivariogram models for comparison to spectral signatures and vegetation indices.

Table 1. Cropping information and image acquisition dates in 1999 and 2000

	Year	Crop	Seeding date	Harvest date	Image acquisition dates (S : compared with Soil properties, Y : compared with Yield data)
Field 1	1999	Corn	May 24	Nov. 5, 6	July 7(Y), August 27(Y)
	2000	Soybean	May 20, 21	Nov. 1, 11	April 12(S), April 26(S), July 25(Y), August 29(Y), September 11(Y)
Gvillo	1999	Soybean	May 10	Oct. 6	July 7(Y), August 27(Y)
	2000	Corn	April 11	Sep. 18, 19	June 25(Y), July 25(Y), September 11(Y)

Fields were grid soil-sampled to a 15 cm depth and analyzed for P (Bray 1 extractable), K, Ca, Mg (ammonium acetate extractable), CEC (sum of bases), organic matter (wet oxidation), salt pH and neutralizable acidity (NA, Woodruff buffer method) using standard University of Missouri procedures (Brown and Rodriguez, 1983). Grid spacing was 33 m for Field 1 and 25 m for Gvillo. Soil sampling point coordinates were later used to extract coincident spectral signatures obtained from the HSI.

Soil EC_a was measured for each field in the fall of 1999 using two commercial sensor systems -- the Geonics EM38 and the Veris 3100. The EM38 operates on the principle of electromagnetic induction and, as operated in the vertical dipole mode, provides an effective measurement depth of approximately 1.5 m. The EM38 was used with a GPS-enabled mobile system described by Sudduth et al. (2001) to collect data every 1 s on measurement transects spaced 10 m apart. The Veris 3100 is a complete commercial system that measures EC_a through coulter electrodes that penetrate the ground surface. This device provides effective measurement depths of approximately 0.3 and 1.0 m. Data was collected every 1 s on a 10 m transect spacing. At the operating speeds used, this time interval corresponded to a 4 to 6 m spacing between sample points. In previous research, we have found these two sensors to provide similar, but not identical mapped EC_a information on claypan-soil fields (Sudduth et al., 1999). EM38 and Veris deep readings have been reliable estimators of claypan-soil topsoil depth (Doolittle, et al., 1994; Kitchen et al., 1999; Sudduth et al., 1999; Sudduth et al., 2001).

3. HYPERSPECTRAL AND MULTISPECTRAL IMAGE ACQUISITION, RECTIFICATION, AND PROCESSING

Airborne images were taken 2 times in 1999 and 7 times in 2000 during the cropping (Table 1). Multispectral images of IKONOS were taken on

June 29, August 4, and September 6, 2000. Soybeans were seeded on May 20 and 21, 2000 for Field 1, which had bare soil at that time due to previous tillage. Field 1 images of HSI taken on April 12 and April 26, 2000 after spring tillage were compared with grid-sampled soil properties and soil EC_a . Since the Gvillo field was cropped under no-tillage and was covered with crop residue and weeds before seeding, no comparison of soil properties and HSI was attempted on that field. Yield data were compared with hyperspectral images taken on July 7 and August 27, 1999 and July 25, August 29, September 11, 2000 for Field 1, and July 7 and August 27, 1999 and June 25, July 25, and September 11, 2000 for Gvillo, and also with multispectral IKONOS images taken on June 29, August 4, and September 6, 2000 both for Field 1 and Gvillo.

The aerial HSI system used in this study was a pushbroom prism-grating scanner (RDACSH3; Real Time Digital Airborne Camera System H3) operated by Spectral Visions Midwest (Mao, 2000). Images were generally acquired between 10:30 a.m and noon. Images included 120 bands from 471 nm to 828 nm (3 nm interval) with a spatial resolution of 1 m and 0.0015 μ m Full Width at Half Maximum (FWHM). Pushbroom scanning is a widely used method for airborne HSI, in which an airborne imaging sensor acquires one image line at a time while the aircraft provides a mobile platform to carry the sensor across the target area.

Geometric distortion was observed in many images, probably due to aircraft attitude change during image acquisition. In general, such geometric distortion should be corrected at the system acquisition level. We applied a rubber sheeting model, which uses piecewise polynomials for image rectification rather than the linear polynomial transformation. We have very accurate surveyed field boundary vector data and resolution merged IKONOS image with a spatial resolution of 1 m taken on August 4, 2000 for both fields. IKONOS image was registered and matched with field boundary and used as a reference image for georeferencing airborne imagery. Rubber sheeting

models are not recommended for rectification of area outside field of interest because of geometric uncertainty, and should be used only when the geometric distortion is severe, ground control points are abundant and no other geometric model is applicable (ERDAS Field Guide, 1997). Foghani (2000) reported that a more precise image was obtained by using a rubber sheeting procedure, compared to polynomial adjustment or an orthophotography algorithm. All HSI data were resampled with a spatial resolution of 4 m due to the file size and the purpose to compare with IKONOS, and converted to ascii format for statistical data set preparation. Landsat-like bands set was created by integrating spectral reflectance values of HSI to the real TM bands, which are the same band range as IKONOS and used for data analysis.

Atmospheric and bidirectional reflectance distribution factor (BRDF) effects should be considered to compensate for solar angle, elevation effects and seasonal change. In this study, chemically-treated reference tarps with known reflectances are laid out for aerial image normalization during flight for two images out of 18 images and then percent reflectance were calculated using the regression model of pixel values of HIS for known reflectance of tarps. Two images with tarps taken on April 12 and 26, 2000 were calibrated radiometrically and used in comparison with ground sensed soil properties such as soil EC_a and chemical properties. For this study, hyperspectral images were generally acquired under clear skies and at a constant low altitude of approximately 1200 meters to get the specified 1 m pixel size imagery. Therefore compensation for atmospheric effects was not applied and digital numbers (DNs) of each image were used for statistical analysis in this study. The BRDF effect may influence pixel values more than atmospheric effect does in aerial imaging. The impact of BRDF distribution functions is still not well understood despite the fact that we know it exists in much of the commonly used remotely sensed (Jensen, 2000).

Vegetation indices are defined as dimensionless, radiometric measures that function as indicators of

relative abundance and activity of green vegetation. These indices could be related as functions of leaf area index (LAI), percentage green cover, chlorophyll content, green biomass, and absorbed photosynthetically active radiation (APAR) (Jensen, 2000; Hong, 1999; Thenkabail et al., 2000; Wiegand et al., 1991). A vegetation index should maximize sensitivity to plant biophysical parameters, normalize or model external effects such as sun angle, viewing angle, and the atmosphere, normalize internal effects such as canopy background variations, and be coupled to some specific measurable biophysical parameter such as biomass, LAI, or APAR (Jensen, 2000). A sensitivity analysis assessing wavelength combinations on spectral indices for explaining yield variability was performed. 10 bands in each of the near infrared (0.76 μm ~0.79 μm), red (0.64 μm ~0.67 μm), green (0.52 μm ~0.55 μm), and blue (0.47 μm ~0.50 μm) wavelengths, whose range were stable for correlation coefficient continuity, were selected for all possible bi-band combinations for the following two indices calculation. All possible bi-band ratio vegetation indices ($RVI = NIR / (RED \text{ or } GREEN \text{ or } BLUE)$) and normalized difference vegetation index ($NDVI = (NIR - RED \text{ or } GREEN \text{ or } BLUE) / (NIR + RED \text{ or } GREEN \text{ or } BLUE)$) of HSI, 300 combinations of wavelengths for calculating the indices, were computed and then correlated with yield data.

4. DATA ANALYSIS

Principal component analysis (PCA) was completed on each image and used as a data set for further statistical analysis. PCA is a procedure for transforming a set of correlated variables into a new set of uncorrelated variables, the principal components (PCs). This transformation is a rotation of the original axes to new orientations that are orthogonal to each other, thus there is no correlation among the transformed variables. Another property of PCA is that the majority of the information contained in a large set of highly correlated variables (wavelengths, in this case) can be represented with a much smaller number of

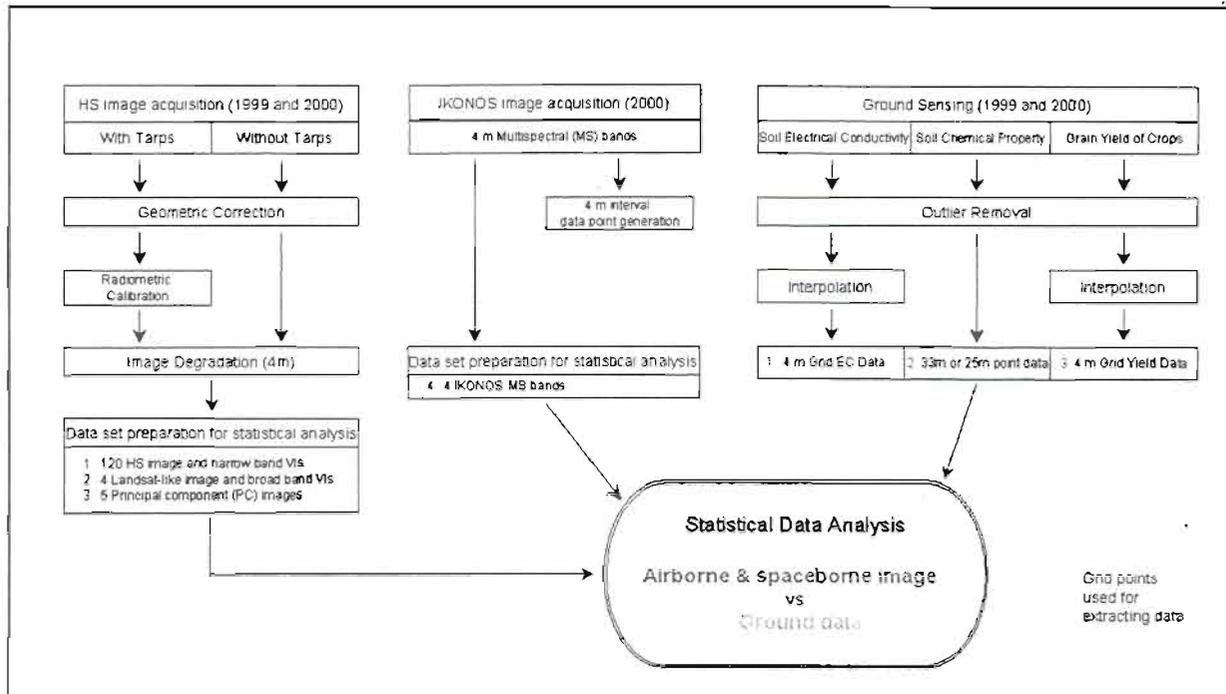


Fig. 1. The flow chart of procedures in this study.

PCs. The first five (PCs) of each image were calculated and used for data analysis. Standard correlation and stepwise multiple linear regression (SMLR) analyses were carried out to determine the relationship between image signatures and ground-collected soil and crop yield. Using the GPS coordinates for soil EC_a, soil chemical properties, and yield, pixel values of coincident points on the imagery were extracted, and PCs calculated. Using SMLR, soil EC_a and chemical property data were regressed against image data consisting of 120 bands, 4 Landsat-like bands, 4 IKONOS bands, and 5 PCs of the Field 1 images taken in April 12 and 26, 2000. Similarly, yield data were regressed against hyperspectral image signatures (120 bands, 4 Landsat-like bands, 4 IKONOS bands, and 5 PCs of the Field 1 images taken in April 12 and 26, 2000. Similarly, yield data were regressed against hyperspectral image signatures (120 bands, 4 Landsat-like bands, 4 IKONOS bands, and 5 PCs) from several Field 1 and Gvillo images (Table 1). The flow chart of procedures are shown in Fig. 1.

5. RESULTS AND DISCUSSION

A. Soil Properties and Hyperspectral Signatures

The spectral reflectance of soil is influenced by moisture content, organic matter, particle size, iron oxide, mineral composition, soluble salts, parent materials, and other factors (Baumgardner et al., 1985). It has also been noted that the environmental conditions under which soils have been formed affect soil reflectance. To investigate the specific relationships present in this data, correlation analysis was completed and correlation coefficients (*r*) were plotted against wavelength (Fig. 2 and 4) to investigate the effective wavelength range as related to soil EC_a and chemical properties - pH, NA, organic matter, P, Ca, Mg, K, and CEC. HSI taken on April 12 and 26, 2000 used for the analysis were calibrated using known-reflectance tarps.

Soil EC_a had a strong negative correlation with all 120 bands and Landsat-like bands (LBs) (Fig. 2). The blue wavelengths showed the highest correlation with EC_a, with the correlation decreasing rapidly as the wavelength increased to around 0.56 μ m. In the green and red wavelengths, correlation coefficients for EC_a were essentially

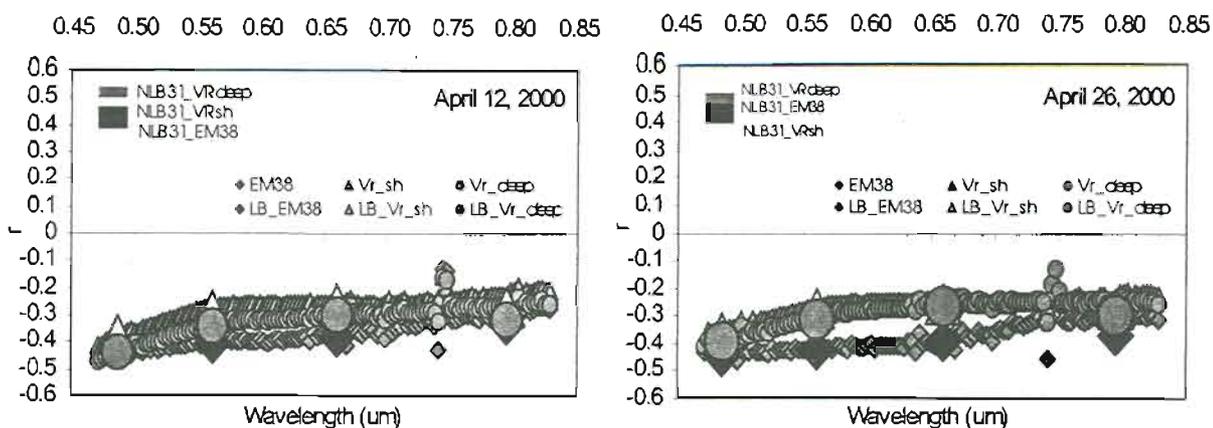
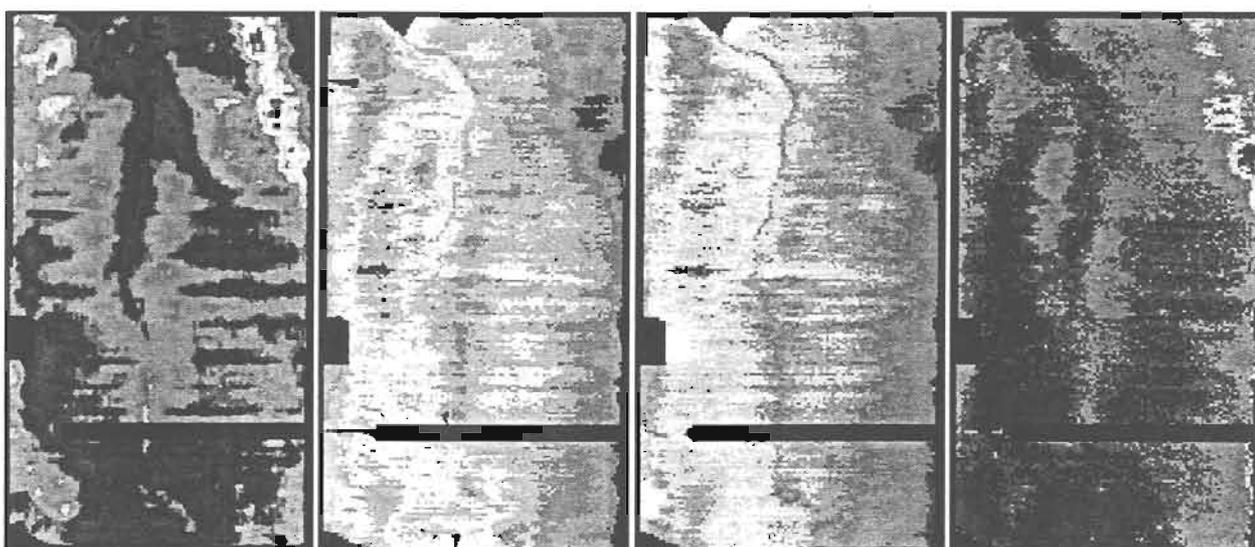


Fig. 2. Correlations of 120 wavelengths, Landsat-like bands(LB), and NLB31($(LB3-LB1)/(LB3+LB1)$) to EC readings obtained by EM38 and Veris Systems.



a) Veris deep sensing data b) Landsat-like band 1 c) First principal component d) $NLB31=(LB3-LB1)/(LB3+LB1)$
 Fig. 3. Soil EC map and images derived from HS data of April 2000.

constant. The Veris 3100 deep (0-1.0 m) EC_a reading was most highly correlated with a spectral index, NLB31 $(LB3-LB1)/(LB3+LB1)$. Around 0.74-0.75 μm a region of noisy data was found in each correlogram (Fig. 2). This is the location of an O_2 and H_2O absorption band, where radiant energy is absorbed by these atmospheric constituents (Jensen, 2000).

SMLR analysis was applied to the soil EC_a and hyperspectral bands for the April 2000 images from Field 1. The most predictive models with coefficients of determination (R^2) ranging from 0.297 to 0.388, are shown in Table 2. The model using EC_a data measured by Veris 3100 deep

sensing was the best significant model for Field 1 with $R^2=0.388$. In addition to the full model, a conservative SMLR model, with $R^2 = 95\%$ of the full model was determined (Table 2). The intention of this model was to reduce the chance of overfitting the data as compared to the full model. A similar approach worked well in a previous spectral data analysis (Sudduth and Hummel, 1991). For all models (Table 2), the number of wavelengths used was reduced by this approach over 50% over the full model, suggesting that little information was contained in those additional data. Five PCs derived from the 120 hyperspectral bands were correlated with soil EC_a data (Table 3). All

five PCs were significantly related to EC_a readings. salinity, organic compounds, and metals (Geonics

Table 2. Stepwise multiple regression model for estimating soil electric conductivity from image data

Dependent Variable	Image Date	Full stepwise Model		R ²	N. of λ s	Conservative Model*							
		R ²	N. of λ s			λ s (um)							
EC-EM38	Apr.12	0.310***	74	0.295***	12	0.471, 0.474, 0.477, 0.498, 501, 0.669, 0.672, 0.687, 0.702, 0.738, 0.741, 0.744							
	Apr.26	0.322***	66	0.306***	10	0.471, 0.480, 0.498, 0.501, 0.669, 0.672, 0.732, 0.741, 0.744, 0.747							
EC-VRsh	Apr.12	0.340***	80	0.323***	30	0.471, 0.474, 0.477, 0.480, 0.483, 0.486, 0.489, 0.504, 0.531, 0.540, 0.543, 0.546, 0.549, 0.552, 0.555, 0.558, 0.561, 0.567, 0.570, 0.573, 0.579, 0.588, 0.654, 0.681, 0.705, 0.708, 0.723, 0.738, 0.741, 0.744							
	Apr.26	0.297***	81	0.282***	36	0.471, 0.474, 0.477, 0.480, 0.483, 0.486, 0.489, 0.492, 0.495, 0.525, 0.531, 0.534, 0.540, 0.543, 0.546, 0.552, 0.555, 0.558, 0.561, 0.570, 0.573, 0.579, 0.588, 0.591, 0.654, 0.669, 0.678, 0.684, 0.699, 0.711, 0.741, 0.744, 0.792, 0.795, 0.807, 0.813							
EC-VRdeep	Apr.12	0.388***	71	0.369***	18	0.471, 0.474, 0.477, 0.480, 0.483, 0.489, 0.495, 0.498, 0.561, 0.567, 0.570, 0.573, 0.576, 0.579, 0.588, 0.672, 0.741, 0.744							
	Apr.26	0.330***	75	0.313***	25	0.471, 0.474, 0.477, 0.480, 0.483, 0.489, 0.495, 0.498, 0.552, 0.558, 0.561, 0.567, 0.570, 0.579, 0.588, 0.591, 0.603, 0.642, 0.669, 0.741, 0.744, 0.747, 0.792, 0.807, 0.828							

* Minimum model which yielded R² ≈ 95% of full model R²

PC1 of the HS image was highly, negatively correlated with soil EC_a and explained over 80% of the variance in the HS data. Multiple regression of the PC data (Table 3) was much less predictive of soil EC_a than SMLR of the original data (Table 2). Within-field EC_a variability map was made using the LB1, 1st principal component, and a spectral index (NLB31) of the imagery taken on April 12, 2000 (Fig. 3 (b), (c), (d)). Soil EC_a can be affected by a number of different soil properties including clay content, soil water content (Kachanoski et al., 1990), varying depths of conductive soil layers (Doolittle et al., 1994), temperature,

Limited, 1992). On these claypan soil fields, soil EC_a is usually highest on eroded side-slopes. Here the claypan is often exposed and therefore the surface will have much higher clay content than at other landscape positions. We have also found soil organic matter higher on eroded side-slopes than in other landscape positions (data not included). We hypothesize that variation in soil clay content (primary factor) and organic matter (secondary factor) are the major soil conditions that provide this relationship between spectral information and soil EC_a.

Soil chemical properties were related to blue.

Table 3. Relationships between soil EC readings and image-derived principal components

Date	VAR ^a		PC1		PC2	
	Apr.12	Apr.26	Apr.12	Apr.26	Apr.12	Apr.26
	81.36		81.36	82.68	3.90	2.16
Variable	R ² of MR for all PCs		r			
	Apr.12	Apr.26				
EC-EM38	0.175***	0.234***	-0.384***	-0.400***	-0.115***	0.108***
EC-VRsh	0.082***	0.090***	-0.283***	-0.280***	-0.034***	0.016*
EC-VRdeep	0.125***	0.133***	-0.337***	-0.304***	-0.045***	0.025
Variable	PC3		PC4		PC5	
	Apr.12	Apr.26	Apr.12	Apr.26	Apr.12	Apr.26
	0.92	0.94	0.56	0.66	0.38	0.40
Variable	r					
EC-EM38	0.113***	-0.144***	0.045***	-0.235***	0.040***	0.014
EC-VRsh	0.029***	-0.060***	0.017*	-0.096***	-0.009	0.007
EC-VRdeep	0.091***	-0.110***	0.047***	-0.175***	0.020**	0.017*

* = 0.05 > p > 0.01, ** = 0.01 > p > 0.001, *** = p < 0.001, a % variance explained by each PC.

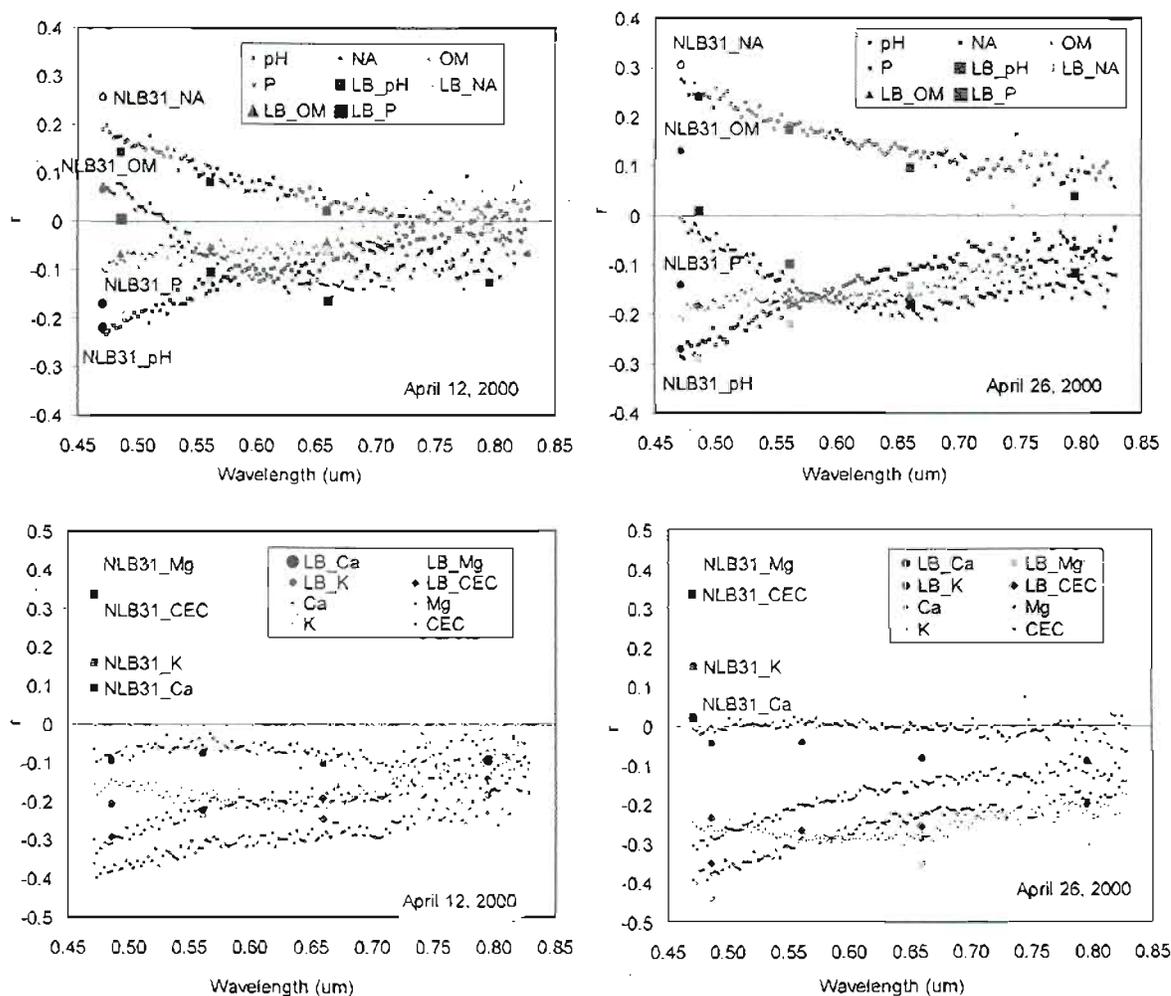


Fig. 4. Correlations of 120 wavelengths, Landsat-like bands(LB), and NLB31((LB3-LB1)/(LB3+LB1)) to soil chemical properties.

properties were determined by soil color and thus, the factors influencing soil color also influenced soil chemical property variability. Except pH, all soil fertility data were negatively correlated to hyperspectral data. Correlations of Mg and CEC to hyperspectral bands and LBs were the highest over the range of wavelengths. For all soil properties, the highest correlations were generally found from 0.47 μm to 0.52 μm , in the blue bands of the visible region, and with NLB31. SMLR analysis was applied to the soil chemical properties and hyperspectral images obtained on April of 2000 for Field 1. The best significant models for soil chemical properties, with R^2 values ranging from 0.270 to 0.416 were obtained on April 12 (Table 4). The models for NA, P, and CEC were the most predictive, with R^2 values of 0.350, 0.415, and 0.416, respectively. Compared to EC_a models, models for grid-sampled soil fertility properties included fewer wavelengths and the reduction in wavelength number from the full to conservative model was less (Table 4). This may have been due to the much smaller number of observations available for the soil fertility data ($n = 365$) as compared to the EC_a data ($n > 8900$).

The five PCs derived from 120 hyperspectral bands were correlated with soil chemical property data (Table 5). PC1 of the HS image explained over 80% of the variance in the HS data. The highest correlations were found between PC1 and Mg, K, and CEC, and between PC2 and pH, NA, and OM. PC-based multiple regression was less predictive of soil fertility data than was SMLR based on the original wavelength data.

B. Yield Data and Hyperspectral Signatures

Hyperspectral image signatures in the 120 individual narrow bands, Landsat-like bands (LBs), IKONOS bands (IBs) were correlated with combine yield data for Field 1 and Gvillo in 1999 and 2000 (Fig. 5). Corn was seeded on May 24 and harvested on November 5 and 6 in 1999 at Field

1. Hyperspectral image signatures taken on July 7, 1999 were poorly correlated to yield data. At this stage, 43 days after seeding (DAS), the crop was at about V11 and some soil was still visible through the corn canopy. About one month after flowering, the August 27 date, the hyperspectral signatures were correlated with yield data, negatively in the visible and positively in the near infrared region. This response is expected since wavelengths in the visible region are closely related to light absorption by plant pigments such as chlorophylls. On the other hand, plants do not absorb near-infrared light and thus reflect more in the near infrared region. Correlation coefficients (r) of LBs for yield were almost the same as those for the corresponding HS wavelengths for each image date except for July 25, 2000. In 2000, corn was seeded on April 11 and harvested on September 18 and 19 at Gvillo. On June 25 the corn crop was in the early reproductive stage (74 DAS), and still green enough to show a typical correlogram between yield and the spectral signature. As corn plants matured, the spectral signature was not correlated with the final yield in the near infrared region on July 25, but a high correlation was still found in the visible region. LB correlations were higher than HS correlations in the visible region. On September 11, one week before harvest, the corn canopy had senesced. The correlogram did not show any distinct trends except in near infrared region, where yield data and the spectral signature were inversely related. In both years, the best relationships between final yield data and spectral signatures were found in the early reproductive stage. On August 4, 2000 (IKONOS), the corn crop (Gvillo) was in the early reproductive stage (85 DAS), and still green enough to show the highest correlation between yield and spectral signature. There was no imagery available during the late vegetative stage, but signatures at that stage would also be expected to show high correlations to yield. Similarly, Hong et al.(1997) reported reflectance of red wavelength was highly related with final yield in booting stage and band ratio (NIR/GREEN) in heading stage for rice

Table 4. Stepwise multiple regression model for estimating soil fertility from image data

Variable	Image Date	Full stepwise Model		Conservative Model*		
		R ²	No. of λ s	R ²	No. of λ s	λ s (μ m)
pH	Apr.12	0.337***	18	0.320***	14	0.474, 0.510, 0.531, 0.546, 0.567, 0.618, 0.687, 0.717, 0.723, 0.726, 0.753, 0.777, 0.828
	Apr.26	0.358***	20	0.340***	16	0.471, 0.507, 0.540, 0.549, 0.555, 0.603, 0.684, 0.690, 0.708, 0.711, 0.714, 0.726, 0.747, 0.753, 0.789, 0.795
NA	Apr.12	0.350***	17	0.332***	14	0.474, 0.498, 0.522, 0.525, 0.537, 0.567, 0.618, 0.687, 0.717, 0.726, 0.741, 0.753, 0.777, 0.825
	Apr.26	0.412***	21	0.391***	16	0.471, 0.474, 0.507, 0.540, 0.549, 0.555, 0.600, 0.684, 0.708, 0.726, 0.753, 0.762, 0.783, 0.789, 0.795, 0.807
OM	Apr.12	0.287***	24	0.272***	22	0.480, 0.507, 0.531, 0.537, 0.543, 0.576, 0.627, 0.666, 0.675, 0.678, 0.714, 0.720, 0.726, 0.735, 0.741, 0.744, 0.747, 0.759, 0.768, 0.792, 0.801, 0.822
	Apr.26	0.321***	21	0.304***	17	0.474, 0.531, 0.546, 0.549, 0.594, 0.645, 0.657, 0.666, 0.687, 0.717, 0.729, 0.744, 0.762, 0.783, 0.810, 0.816, 0.828
P	Apr.12	0.415***	21	0.394***	16	0.486, 0.492, 0.495, 0.516, 0.531, 0.537, 0.627, 0.645, 0.654, 0.663, 0.708, 0.726, 0.735, 0.774, 0.798, 0.822
	Apr.26	0.420***	27	0.399***	21	0.471, 0.483, 0.492, 0.534, 0.543, 0.555, 0.561, 0.576, 0.615, 0.621, 0.627, 0.645, 0.648, 0.654, 0.681, 0.696, 0.702, 0.714, 0.741, 0.762, 0.798
Ca	Apr.12	0.270***	23	0.256***	21	0.477, 0.504, 0.531, 0.534, 0.543, 0.552, 0.567, 0.570, 0.612, 0.642, 0.666, 0.699, 0.702, 0.711, 0.714, 0.723, 0.747, 0.759, 0.765, 0.774, 0.789
	Apr.26	-				
Mg	Apr.12	0.334***	18	0.317***	12	0.474, 0.486, 0.495, 0.525, 0.543, 0.564, 0.621, 0.741, 0.753, 0.783, 0.798, 0.801
	Apr.26	0.415***	26	0.394***	20	0.471, 0.483, 0.507, 0.543, 0.561, 0.603, 0.612, 0.615, 0.624, 0.639, 0.675, 0.678, 0.696, 0.705, 0.744, 0.753, 0.783, 0.789, 0.801, 0.825
K	Apr.12	0.279***	22	0.265***	19	0.474, 0.504, 0.516, 0.537, 0.627, 0.654, 0.657, 0.660, 0.708, 0.723, 0.726, 0.735, 0.753, 0.762, 0.774, 0.783, 0.804, 0.819, 0.822
	Apr.26	0.304***	13	0.289***	10	0.501, 0.558, 0.594, 0.645, 0.681, 0.696, 0.714, 0.717, 0.726, 0.744
CEC	Apr.12	0.416***	33	0.395***	28	0.474, 0.477, 0.486, 0.495, 0.516, 0.531, 0.537, 0.543, 0.570, 0.594, 0.603, 0.618, 0.630, 0.654, 0.699, 0.702, 0.711, 0.714, 0.717, 0.726, 0.738, 0.741, 0.753, 0.783, 0.792, 0.798, 0.822
	Apr.26	0.382***	23	0.363***	19	0.474, 0.483, 0.507, 0.573, 0.603, 0.612, 0.615, 0.624, 0.645, 0.651, 0.666, 0.678, 0.726, 0.729, 0.744, 0.753, 0.783, 0.789, 0.816

* Minimum model which yielded R² \approx 95% of full model R².

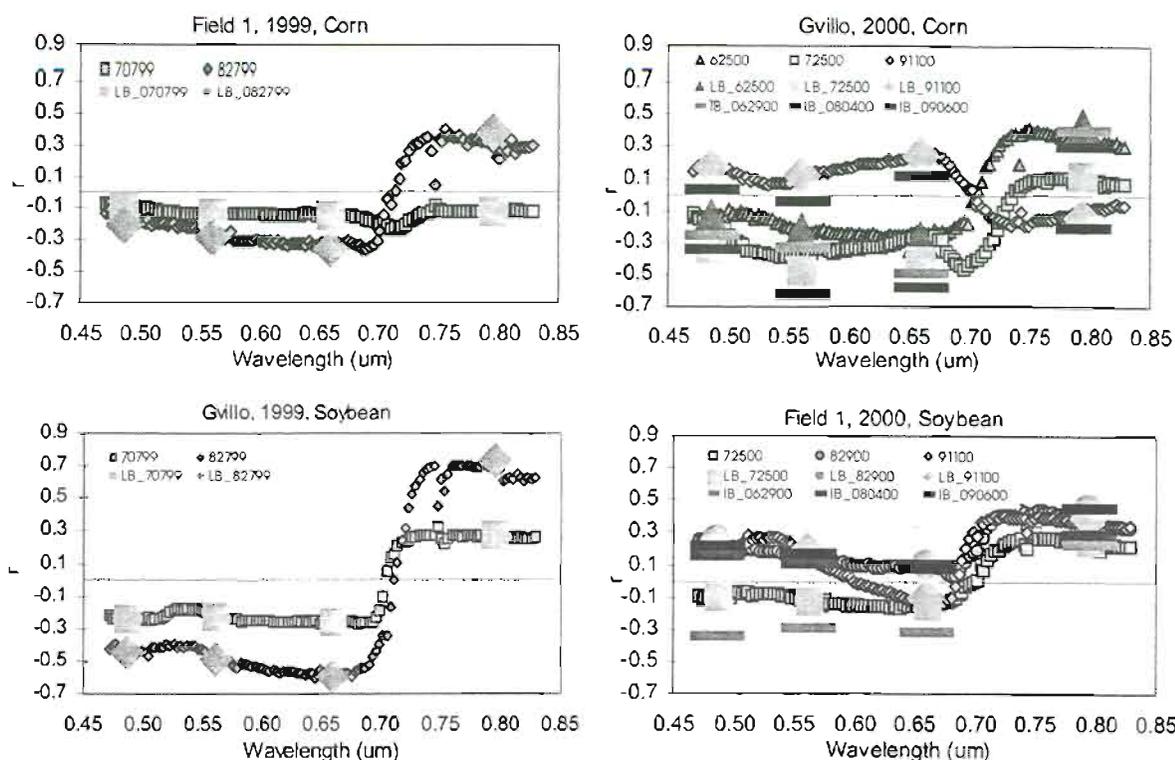


Fig. 5. Correlations of 120 wavelengths, Landsat-like bands(LBs), and IKONOS Bands(IBs) to yield data.

canopy.

For soybeans, yield data were generally most highly correlated with wavelengths in the near infrared region (Fig. 5). About 100 DAS soybean plants reflected the most incident near infrared energy in both years. There were no distinct relationships between yield data and spectral signatures in the visible region for any of the images except for the image taken on August 27, 1999. Soybean plants usually flower about 50 DAS depending on weather and crop management conditions. Image acquisition dates for this study were 57 and 107 DAS in 1999 and 65, 99, and 110 DAS in 2000. Although spectral signatures of corn plants were obvious in the visible and near infrared region at different growth stages, signatures of soybean did not change appreciably after flowering. We believe these signature differences to be caused by differences in growth habit between the two crops, especially during the reproductive stage. Soybean is indeterminate and corn is determinant in flowering behavior. It usually takes 5~7 days for flowering in corn and thus affects the

spectral signatures of corn within a short time. On the other hand, soybean flowers over five to six weeks. Soybean flowers are behind broad-leaves, which remain green for a long time, and spectral signatures are not influenced a lot by flowering.

Ten narrow bands for each wavelength range NIR(0.76~0.79), RED(0.64~0.67), GREEN(0.52~0.55), and BLUE(0.47~0.50) – were selected based on the correlogram (Fig. 5). The correlations with yield of each of the 300 RVI and NDVI indices obtained for each image date were expressed as a scatterplot (Fig. 6). For corn, RVI and NDVI correlations were very similar in both years. Correlations between yield data and RVI and NDVI of August 27, 1999 were the highest obtained. In 2000, high correlations were found when using the RVI and NDVI data of June 25. Vegetation indices did not improve correlation coefficients with yield, compared to raw HS data and none of the individual band combinations provided significantly better correlations than any other combinations within each spectral region, implying that there was no advantage in this case

Table 5. Relationships between soil fertility and image-derived principal components

Variable	Date		PC1		PC2	
	R ² of MR for all PCs		Apr.12	Apr.26	Apr.12	Apr.26
	Apr.12	Apr.26	84.48	84.11	1.40	1.92
pH	0.095***	0.156***	0.023	0.096	0.289***	-0.383***
NA	0.123***	0.197***	-0.056	-0.140**	-0.338***	0.412***
OM	0.114***	0.148***	-0.067	-0.122*	-0.333***	0.355***
P	0.022	0.018	-0.125*	-0.122*	-0.064	-0.016
Ca	0.017	0.019	-0.095	-0.077	0.016	-0.081
Mg	0.144***	0.202***	-0.267***	-0.351***	-0.283***	0.321***
K	0.105***	0.126***	-0.213***	-0.239***	-0.256***	0.266***
CEC	0.117***	0.150***	-0.181***	-0.249***	-0.303***	0.319***
	PC3		PC4		PC5	
	Apr.12	Apr.26	Apr.12	Apr.26	Apr.12	Apr.26
	0.75	0.90	0.60	0.78	0.48	0.49
Variable	r					
pH	0.081	-0.010	-0.047	0.077	-0.030	-0.066
NA	-0.061	0.022	0.030	-0.132*	0.029	0.074
OM	-0.056	0.060	-0.043	-0.108*	-0.077	0.089
P	-0.036	0.044	-0.047	-0.014	-0.054	0.001
Ca	0.007	0.014	-0.070	0.063	-0.038	0.007
Mg	-0.055	0.051	-0.079	-0.011	-0.023	0.014
K	-0.002	0.016	-0.044	-0.010	0.003	0.064
CEC	-0.051	0.038	-0.049	-0.054	-0.012	0.061

* = 0.05 > p > 0.01, ** = 0.01 > p > 0.001, *** = p < 0.001, a % variance explained by each PC.

for hyperspectral data as compared to multispectral images.

In soybean, correlations with RVI and NDVI were essentially the same. High correlations between yield data and RVI and NDVI were obtained with the August 27, 1999, when was 107 DAS. There was no apparent difference in correlation coefficient for the different narrow band combinations in soybean as well as in corn. Vegetation indices itself did not improve correlation coefficients with yield. The best time to show high

correlation between yield and vegetation indices coincided with the time for spectral signatures. Multiple regression analysis using temporal data is needed for improved explanation of yield data.

Relationships between yield and the first 5 PCs of image data, along with the R² of multiple regression models for estimating crop yield using all five PCs are shown in Table 6. All five PCs derived from 120 HS bands were correlated with yield data both for corn and soybean. The best model explaining yield using 5 PCs was obtained

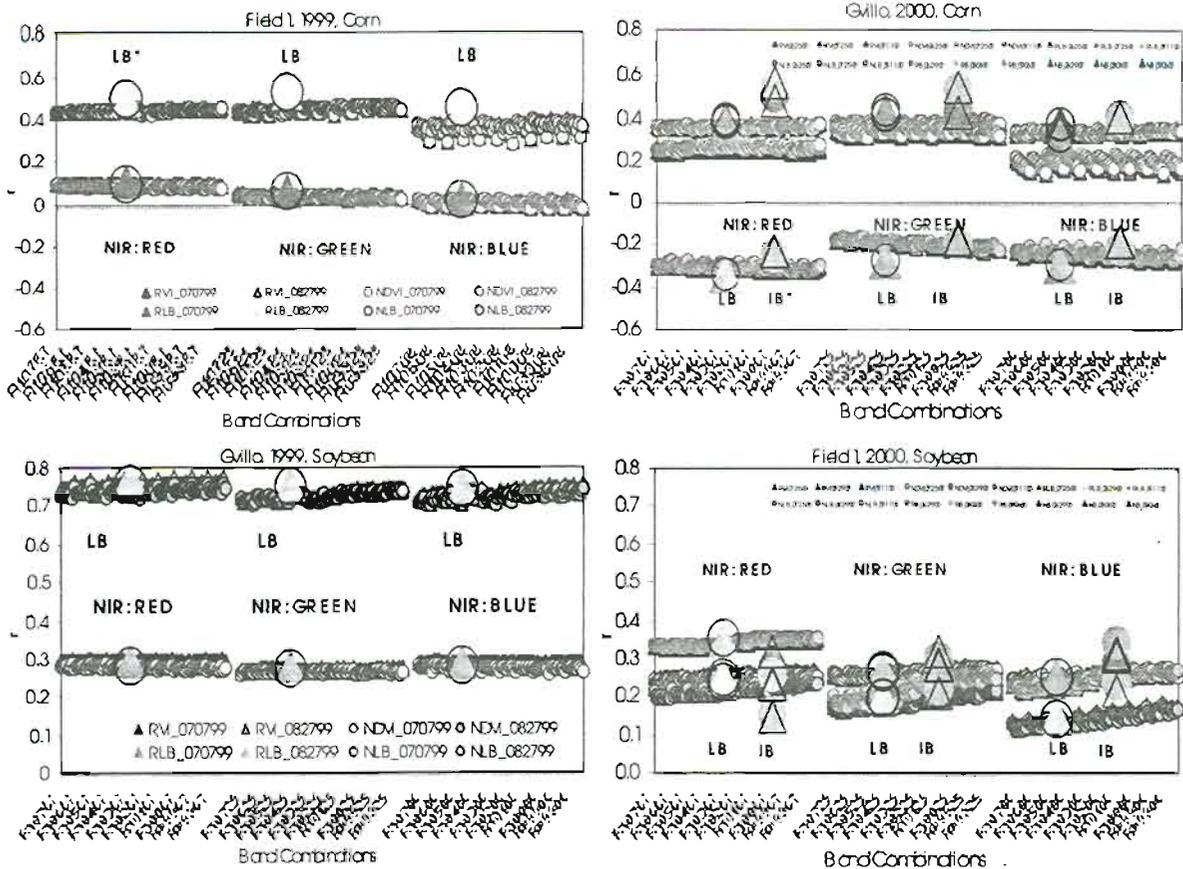
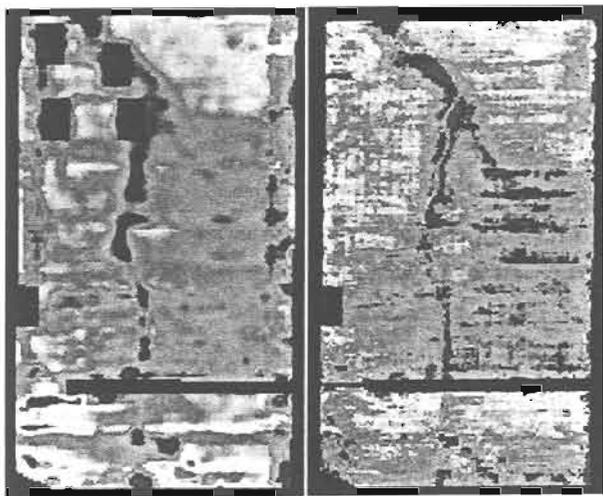


Fig. 6. Correlations of ratio vegetation index (RVI) and normalized difference vegetation index (NDVI) from 120 wavelengths, Landsat-like bands(LBs), and IKONOS bands(IBs) with yield data.



[a] [b]

- a) Grid yield map obtained by yield
- b) IKONOS band 4 (Sep. 6, 2000) by yield sensing system

Fig. 7. Yield map from yield monitoring and IKONOS image for soybean in Field 1.

with the imagery of August 27, 1999 in corn and soybean. In these models, more than 75% and 90% of the variability in final yield for corn and soybean, respectively, were explained by 5 PCs. As the cropping season progressed more PCs were required to explain the variance in the HS data.

A within-field yield variability map made was shown and compared with grid yield map in Fig. 7 and 8. Image of IKONOS band 4 was highly related to the Field 1 soybean yield map from combine yield sensing (Fig. 7). Within-field yield variability map from IKONOS-derived NDVI image taken on August 4, 2000 and multiple linear regressed image with 4 LBs of HS data showed a high spatial relationship with corn yield and soybean yield, respectively, in the Gvillo field (Fig. 8).

Table 6. Relationships between crop yield and image-derived principal components

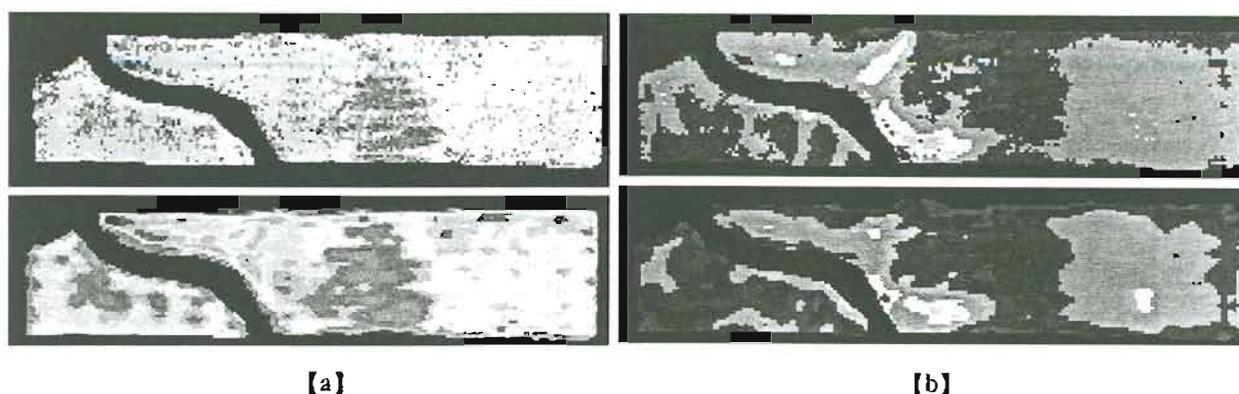
Crop	Field	Year	Image Date	R ² of MR model for all PCs	r				
					PC1		PC2		
Corn	Field 1	1999	Jul. 7	0.189***	-0.138***	64.65	-0.109***	26.90	
			Aug. 27	0.445***	-0.281***	49.99	-0.110***	20.44	
	Gvillo	2000	Jun. 25	0.274***	-0.379***	49.01	0.175***	21.58	
			Jul. 25	0.395***	-0.377***	35.58	0.407***	18.26	
			Sep. 11	0.162***	-0.157***	57.07	-0.314***	15.82	
Crop	Field	Year	Image Date	R ² of MR model for all PCs	r				
					PC3		PC4		PC5
Corn	Field 1	1999	Jul. 7	-0.180***	0.85	-0.320***	0.73	0.081	0.31
			Aug. 27	0.194***	5.08	-0.051***	1.09	-0.170***	1.01
	Gvillo	2000	Jun. 25	0.130***	2.51	-0.311***	1.87	-0.434***	0.90
			Jul. 25	-0.458***	2.23	-0.023*	1.80	0.111***	1.08
			Sep. 11	-0.168***	2.28	0.167***	0.60	-0.193***	0.53
Crop	Field	Year	Image Date	R ² of MR model for all PCs	r				
					PC1		PC2		
Soybean	Gvillo	1999	Jul. 7	0.126***	-0.270***	79.41	0.082***	13.04	
			Aug. 27	0.673***	-0.547***	62.15	0.591***	24.91	
	Field 1	2000	Jun. 25	0.142***	0.182***	56.93	-0.221***	32.17	
			Jul. 25	0.292***	0.430***	47.10	0.163***	24.36	
			Sep. 11	0.195***	0.304***	54.73	0.254***	28.05	
Crop	Field	Year	Image Date	R ² of MR model for all PCs	r				
					PC3		PC4		PC5
Soybean	Gvillo	1999	Jul. 7	-0.046***	0.72	0.027*	0.46	0.207***	0.42
			Aug. 27	0.031***	1.72	-0.144***	0.70	0.019	0.56
	Field 1	2000	Jun. 25	-0.016*	1.45	-0.187***	0.88	0.058***	0.48
			Jul. 25	0.136***	1.94	-0.153***	1.05	0.014	0.83
			Sep. 11	-0.090***	2.36	0.105***	0.95	0.027***	0.80

* = 0.05 > p > 0.01, ** = 0.01 > p > 0.001, *** = p < 0.001, a % variance explained by each PC.

6. CONCLUSIONS

Several statistical methods - correlation analysis, SMLR, and PCA were successfully used to relate within-field information on soils and crops with hyperspectral imagery, Landsat-like bands, IKONOS imagery. Hyperspectral image signatures of bare soil taken on April, 2000 were highly correlated with soil EC_a and chemical properties. Blue wave-

lengths in the visible region, Landsat-like band 1, and the 1st PC of HS data were informative for soil EC_a. Soil chemical properties were related to blue, green, and red wavelengths in the visible region rather than wavelengths in the near infrared region, which implied that spectral reflectance signatures from soil surface were usually determined by soil colors. Thus, the factor influencing soil colors also had an important role to represent



a) NDVI_IKONOS image of Aug. 4, 2000 (top) and grid yield map obtained by yield sensing system (bottom) map (bottom)
 b) Multiple linear regression model of 4 LBs for yield estimation from HS image of Aug. 27, 1999 (top) and grid yield map (bottom)

Fig. 8. Yield maps from yield monitoring and airborne and satellite images for Gvillo in 1999 and 2000.

the spectral reflectance signatures. Soil moisture undoubtedly played an important role. Within-field ECa variability map was derived from LB1, 1st PC, and a spectral index (NLB31) of the imagery taken on April 12, 2000.

Hyperspectral image signatures in the 120 individual narrow bands, Landsat-like bands, principal components, and vegetation indices (RVI and NDVI) derived from hyperspectral images, and IKONOS image signatures were highly and significantly correlated with final yield of corn and soybean if acquired at the proper growth stage. Highest correlations to corn yield were generally found in the visible region, while highest correlations to soybean grain yield were generally found in the near infrared region. Highest correlations between yield data and RVI and NDVI were obtained with the August 27, 1999 data for both corn and soybean. Correlations with RVI and NDVI were essentially the same. Within-field soybean yield map was made using IKONOS band 4 from the imagery of September 6 in Field 1. Within-field yield variability map from IKONOS-derived NDVI image taken on August 4, 2000 and multiple linear regressed image with 4 LBs of HS data were made for explaining corn yield and soybean yield, respectively, in the Gvillo field.

From PCA, the majority of total variance was explained by the first PC when relating to soil ECa and fertility data. For yield data, as the cropping season progressed more PCs were required to explain the variance in the HS data.

Additionally from this investigation we recommend that the geometric distortion of an airborne pushbroom sensor due to vehicle attitude change should be corrected at the system level. Also, we acknowledge that studying optimal pixel size to know the ratio of signal to random variability or noise ought to consider spatial statistical analysis.

ACKNOWLEDGEMENTS AND DISCLAIMER

We thank the following for financial support or assistance enabling this research: North Central Soybean Research Program, United Soybean Board, Spectral Visions, and USDA-CSREES *Special Water Quality Grants* programs. We also thank Scott Drummond, Clyde Fraise, Jon Fridgen, and Lei Tian for their assistance in analyzing the data. Mention of trade names or commercial products is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the United States Department of Agriculture.

REFERENCES

1. Baumgardner, M. F., L. F. Silva, L. L. Biehl and E. R. Stoner. 1985. Reflectance properties of soils, *Advances in Agronomy*, 38:1-44.
2. Birrell, S. J., K. A. Sudduth and S. C. Borgelt. 1996. Comparison of sensors and techniques for crop yield mapping. *Computers and Electronics in Agriculture*, 14(2/3):215-233.
3. Brown, J. R. and R. R. Rodriguez. Soil testing in Missouri: a guide for conducting soil tests in Missouri. Extension Circular 923. Univ. of Missouri Extension Division, Columbia, Missouri.
4. Doolittle, J. A., K. A. Sudduth, N. R. Kitchen, and S. J. Indorante. 1994. Estimating depths to claypans using electromagnetic induction methods, *J. Soil and Water Cons.*, 49(6):572-575.
5. ERDAS Field Guide. 4th Edition, ERDAS Inc. pp. 158-159, 225, 1997.
6. Forghani A. 2000. Geometric registration of aerial photography using three image correction, In Proceedings of the 2nd International Conference on Geospatial Information in Agriculture and Forestry, Lake Buena Vista, FL, p. 11-531-538 January.
7. Goetz, A., G. Vane, J. E. Solomon and B. N. Rock. 1985. Imaging spectrometry for earth remote sensing, *Science*. 228(4704):1147-1153.
8. Geonics Limited. 1992. Geonics bibliography version 2.2, Oct., 1992. Mississauga, Ontario, Canada.
9. Hong, S. Y. 1997. Radiometric estimates of grain yields related to crop aboveground net production (ANP) in paddy rice. In Proceedings of '97 International Geoscience and Remote Sensing Symposium(IGARSS), Singapore, pp. 1793-1795 August.
10. Hong, S. Y. 1998. Analysis on rice growth information and estimation of paddy field area by using remotely sensed data, Ph.D. Dissertation, Agronomy Department, Kyungpook National University, Daegu, Korea, December.
11. Jensen, J. R. 2000. Remote sensing of the environment: An earth resource perspective, Prentice Hall. Upper Saddle River, New Jersey, p. 42, 227, 347.
12. Kachanoski, R. G., E. DeJong and I. J. Van-Wesenbeeck. 1990. Field scale patterns of soil water storage from non-contacting measurements of bulk electrical conductivity, *Can. J. Soil Sci.* 70:537-541.
13. Kitchen, N. R., K. A. Sudduth and S. T. Drummond. 1999. Soil electrical conductivity as a crop productivity measure for claypan soils. *J. Prod. Agric.* 12:607-617.
14. Mao, C. 2000. Hyperspectral focal plane scanning - An innovative approach to airborne and laboratory pushbroom hyperspectral imaging, In Proceedings of the 2nd International Conference on Geospatial Information in Agriculture and Forestry, Lake Buena Vista, FL, p. 1-424-428 January.
15. Moran, M. S. 2000. Image-based remote sensing for agricultural management - Perspectives of image provider, research scientists and users. In Proceedings of the 2nd International Conference on Geospatial Information in Agriculture and Forestry, Lake Buena Vista, FL, p. 1-23-30 January.
16. Mulla, D. J., A. C. Sekely and M. Beatty, Evaluation of remote sensing and targeted soil sampling for variable rate application of lime, [CD-ROM computer file]. In P.C. Robert et al. (ed.) Proceedings of 5th International Conference on Precision Agriculture, Minn., MN, Jul. 16-19, 2000. ASA, CSSA, and SSSA, Madison, WI.
17. Sudduth, K. A. and J. W. Hummel. 1991. Evaluation of reflectance methods for soil organic matter sensing, *Transactions of the ASAE*. 34(4):1900-1909.
18. Sudduth, K. A., S. T. Drummond, S. J. Birrell, and N. R. Kitchen. 1996. Analysis of spatial factors influencing crop yield, In Proceedings of the 3rd International Conference on Precision Agriculture, Minneapolis, MN, p. 129-139 June 1996.
19. Sudduth, K. A., N. R. Kitchen and S. T. Drummond. 1998. Soil conductivity sensing on claypan soils : comparison of electromagnetic

- induction and direct methods. In Proceedings of the 4th International Conference on Precision Agriculture, St. Paul, MN, p. 979-990 July.
20. Sudduth, K. A., N. R. Kitchen and S. T. Drummond. 1998. Soil conductivity sensing on claypan soils: comparison of electromagnetic induction and direct methods. Vol II, p. 979-990. In Proc. 4th International Conference on Precision Agriculture, St. Paul, MN, Jul. 19-22, ASA, CSAA, SSSA, Madison, WI.
21. Sudduth, K. A., S. T. Drummond and N. R. Kitchen. 2001. Accuracy issues in electromagnetic induction sensing of soil electrical conductivity for precision agriculture. *Comp. and Electronics in Agric.* 31:239-264.
22. Thenkabail, P. S., R. B. Smith and E. D. Pauw. 2000. Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote Sensing of Environment*, 71:158-182.
23. Wiegand, C. L., A. J. Richardson, D. E. Escobar and A. H. Gerbermann. 1991. Vegetation indices in crop assessments. *Remote Sensing of Environment*, 35:105-119.
24. Yao, H., A. T. Tian and N. Noguchi. 2001. Hyperspectral imaging system optimization and image processing. In Proceedings of the 2001 ASAE Annual International Meeting, Sacramento, CA, July.